EFFECT OF DATA COVERAGE ON THE ACCURACY OF 500-MB. FORECASTS

CHARLES L. BRISTOR, U. S. WEATHER BUREAU

Joint Numerical Weather Prediction Unit, Suitland, Md.

[Manuscript Received May 22, 1958; Revised July 3, 1958]

ABSTRACT

Networks of data are simulated by interpolating height and wind from a hypothetically "correct" analysis at a uniform array of points. The interpolated values are added to random numbers whose statistics are typical of non-systematic errors of observation, and are then regarded as genuine data. Such artificially constructed "data" for several networks of different densities are analyzed independently. The differences between these analyses and the hypothetically "correct" analysis are taken to be representative of the initial analysis error under conditions of varying station density.

Numerical forecasts computed from the different analyses are compared with the forecast made from the "correct" initial data. Several such comparisons indicate that initial analysis errors do not grow to an important degree so long as the spacing between synoptic reports confines error fields to a scale smaller than that of the synoptic disturbances. However, with data spacing comparable with that over existing regions of poor data coverage, initial errors are amplified two-fold in a 48-hr. forecast interval.

Further experiments, in which the data are analyzed objectively to eliminate inconsistencies of analysis, are carried out.

1. INTRODUCTION

The problem of resolving the initial synoptic situation in all its pertinent detail has often been mentioned as a contributing factor where critical weather developments have challenged the forecaster and found his efforts inadequate. Although forecasters justify requests for more data (at least among themselves) on the basis of this argument, additional data are more often made available in connection with some expanded commercial program such as a new air route where there is direct operational need for current observations of weather conditions. The forecaster must then plead that good forecasts on such routes are also contingent on adequate data coverage in other areas. This plea would perhaps be weighed more heavily except for the subjectivity in the argument. No matter how elegant are the forecast parameters involved or how precisely observational data are employed in their evaluation, the final result is usually a subjective combination which leaves the contribution of the avowed critical factors somewhat in doubt.

This argument could be greatly strengthened if a measure of change in forecast accuracy could be related quantitatively to a change in data coverage.[1] The problem would then be one of economics with the increased cost being weighed against the increased value of a more accurate forecast. Newton [1] has discussed this data and analysis problem at length and has shown how differences in barotropic tendency computations can be attributed to differences in analyses. Best [2] has carried through several barotropic forecasts with similar conclusions. The purpose of the present study is to strengthen this argument in a quantitative way by isolating the numerical forecast errors arising from initial analysis errors that, in turn, are made a function of data density alone. The experiment described below is actually an adjunct to a more theoretical treatment by Thompson [3]. The results, although limited, are considered meaningful in a corroborative way. Computational roundoff and truncation errors remain as contaminants to the results. Other errors such as boundary differences and analysts' subjectivity as discussed later, are controlled insofar as is practical.

[1] Preliminary results of an investigation by Maj. E. O. Jess, USAF, have come to the attention of the writer. His approach is much the same as in the present paper although different measures of forecast error are employed. The two studies agree in a broad sense.

2. DESIGN OF THE EXPERIMENT

The experiment was designed to indicate differences in forecasts (hereafter called errors) arising from differences in initial analyses of the same synoptic situations. Such a definition of error seems entirely fair. The objective was to measure quantitatively the reproduction of barotropic forecasts based on plentiful data by barotropic forecasts based on limited data. The similar task involving the real atmosphere or even more sophisticated models is certain to be more difficult. The barotropic forecasts were produced using an octagonal grid of 1977 points covering most of the Northern Hemisphere (fig. 1).

Four different analyses of two separate synoptic situa-

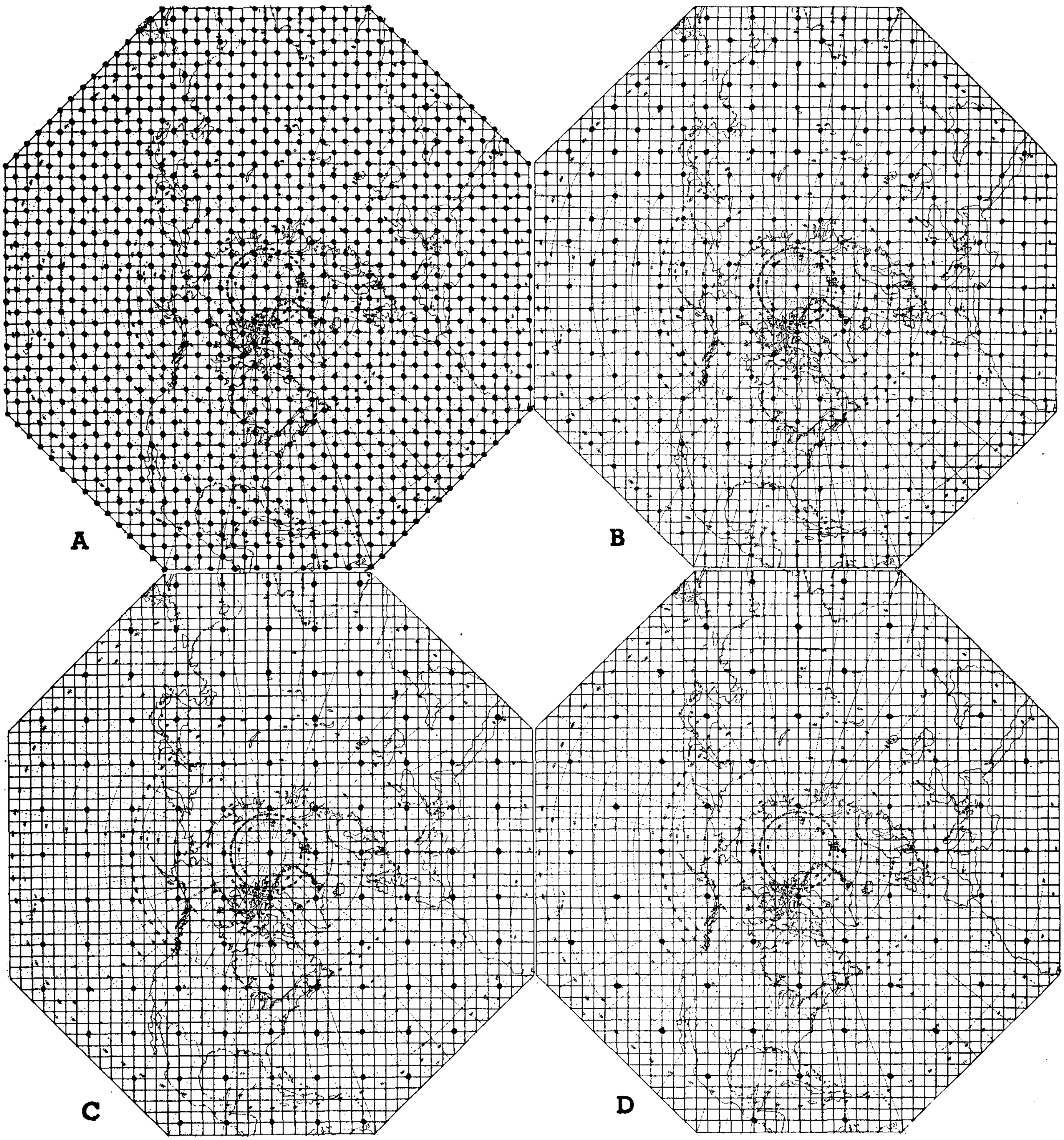

FIGURE 1.—Data coverage, shown by dots, for the octagonal grid of 1977 points currently used in the JNWP Unit (A) and successively reduced arrays (B, C, D,).

tions were provided as input data for the forecasts in the following manner. The 500-mb. data from two routine operational analyses were first interpolated at the grid points indicated by large dots in figure 1A. It was then postulated that these grid-point data were the true heights of the 500-mb. surface at the points in question. The next problem was to try to reproduce these true grid-point values with subjective analyses made under realistic operational conditions but with 500-mb. data coverage being provided in various reduced arrays. The analyst attempting the simulation had access to complete routine operational surface analyses which contained reasonably complete coverage over most geographical regions including ocean acreas. The question was thus reduced to one of upper-air data coverage. Maps for 0300 GMT for April 3 and 5, 1957 were selected partly because data decks for points shown in figure 1A were readily available. Also, for a limited sampling, it is perhaps more meaningful to choose such a period having flow patterns of moderate intensity rather than either winter or summer extremes.

The grid-point 500-mb. heights (considered to be uniformly spaced perfect data) were modified by the application of random normally distributed errors having a mean value of 50 feet. Geostrophic winds measured at the grid points from the basic analysis were likewise amended by similarly applying a non-systematic 10-knot average wind speed adjustment. The realistic but uniformly distributed data thus obtained were next plotted on blank charts in three reduced arrays as shown in figures 1B, C, and D. In some of the illustrations these data arrays are referred to as maximum, intermediate, minimum, and sub-minimum for reasons which will become obvious. The same experienced analyst proceeded to analyze one series, starting with the most sparse data array first. He was provided with two preceding analyses having the same data array and was permitted use of normal differential analysis techniques. Once the coarse array had been analyzed, he was given the denser array and repeated the analysis process.

Figure 2 presents the corresponding final analyses for the data arrays shown in figure 1 for the case of 0300 GMT, April 3, 1957. As one might suspect, the patterns look strikingly similar, at least in a superficial way. Indeed one might say that the differences appear trivial if the chart is to serve as a basis for a subjective forecast. Significant differences do exist, however, and their repercussions in numerical forecasts are important.

3. COMPUTATIONS AND RESULTS

All competing forecasts were made using the operational barotropic forecast procedure being employed at the time of the experiment, except that special precautions were taken to minimize boundary error contamination. In particular an investigative code devised by 'Arnason [4] was employed to remove virtually all boundary inflow-outflow wind components from each set of input data. The technique involves replacing each boundary height by the mean of all boundary values and allows the discrepancy to be "faired" in with a weight of .4 at the first internal ring of points, a weight of .1 at the second ring, and zero correction inward. Such an adjustment may, on occasion, do harm in the already questionable boundary region but, on the basis of other tests, it appears to prevent large effects from penetrating meridionally far into the grid. In any case, rather elusive, but sometimes important, minor inflow-outflow differences have thus been removed at a slight sacrifice in realism which does not enter into the results since the error is defined as the difference between forecasts.

Berggren [5] has recently shown the importance of subjective opinion among analysts in the interpretation of identical plotted data charts. Allowing one analyst to perform all the competing simulation analyses for each case was considered to be the best practical way of alleviating this difficulty. In this way any systematic habits or model concepts held by the analyst should not cause large random elements to enter into the results. The greatest chance for such a discontinuity to enter would be between the given operational analysis and the three simulated analyses but this difference enters more or less equally into all three resulting comparisons. Actually the analyst engaged in the second case was available for only a limited period and was unable to complete the analysis for the data array of figure 1B. The analyst for the first case completed the second case. One would expect this nonhomogeneous effect to be small since the data coverage dealt with was closest to the maximum thereby offering the least opportunity for varying interpretations. The results do not appear over-sensitive to this difficulty but differences between the two cases will be discussed in this light in connection with wind error results. The distribution of geostrophic wind errors was selected as the measure of forecast skill because Thompson's conclusions involve the wind error and also because this measure is perhaps of most interest to those making direct use of 500-mb. prognostic charts. The several mean distances between observations expressed by the data arrays were selected for two reasons: (1) They approximate familiar arrays which presently exist over limited regions, and (2) error fields thus produced have different characteristic wavelengths.

The coverage expressed in figure 1B closely approximates that now in existence over most of Canada. Figure 1C corresponds rather well to the present Atlantic coverage if one excludes reconnaissance data and also excludes a vast area south of 30° N. between Africa and the West Indies. Figure 1C approximates the spacing between present subtropical Pacific islands and corresponds in general to Pacific coverage without reconnaissance. The results thus indicate error levels corresponding to uniform hemisphere-wide data distributions which exist at present over certain segments of the hemisphere. From the standpoint of scale comparisons, the mesh lengths in figure 1B and 1D permit error fields having half wavelengths of approximately 360 n. miles and 720 n. miles, respectively. The half wavelengths of many synoptic disturbances having important local weather anomalies are shorter than the mesh length of figure 1D. There are likewise very few macro-scale disturbances (beyond their embryonic stages) which are characterized by half wavelengths less than the mesh size of figure 1B. The present study thus provides some empirical evidence bearing on the first three items listed by Thompson [3]: (1) Forecast range, (2) Initial wind error, and (3) Difference in scale between disturbances and error fields.

483854—58——2

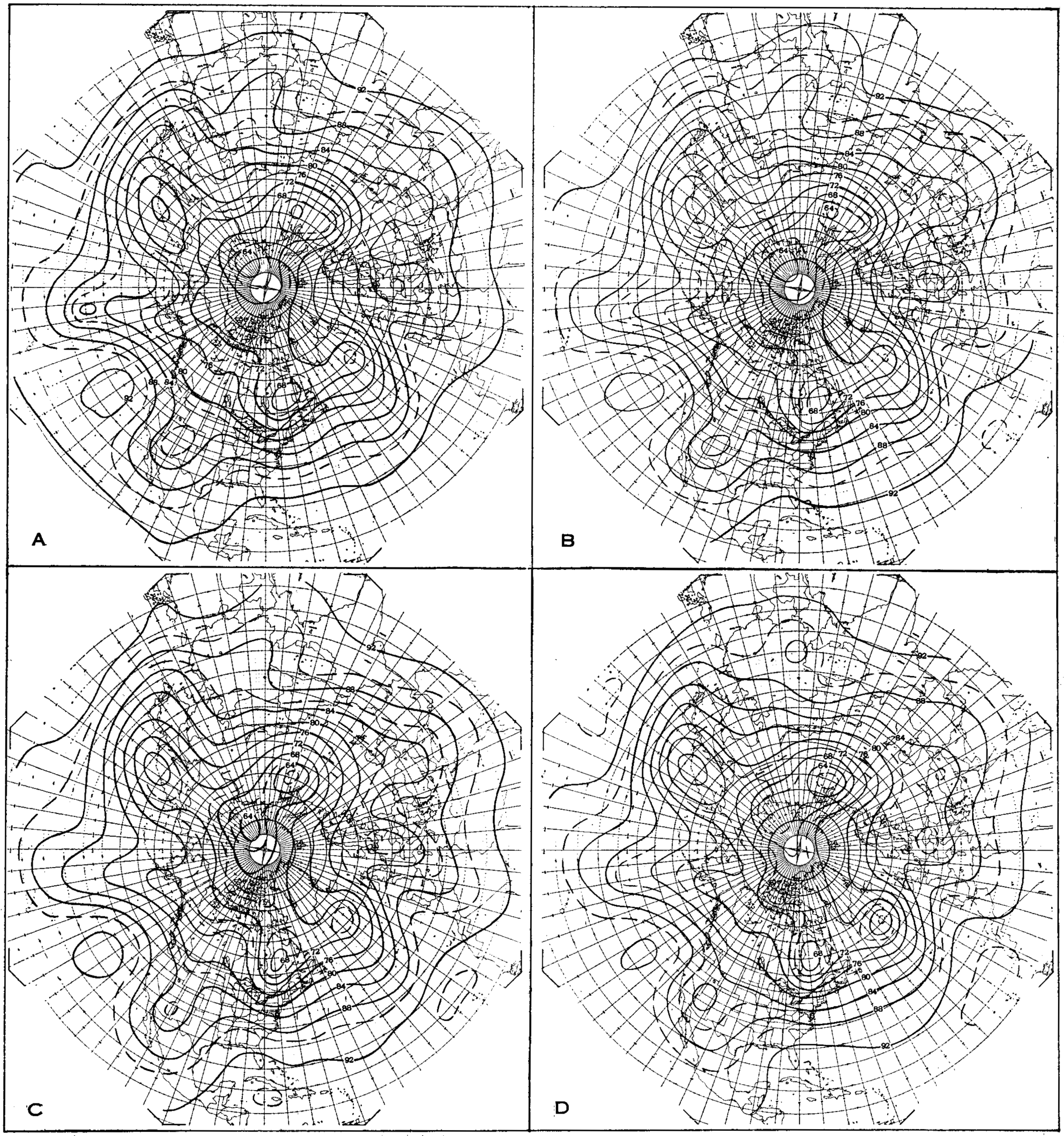

FIGURE 2.—500-mb. analyses for Case 1, 0300 GMT, April 3, 1957 based on data from the grid array with the corresponding letter in figure 1.

To summarize the procedure, the results which follow were obtained from the mentioned analysis procedure through the following computational steps:

a. Boundary heights contained in the input data decks were adjusted to a constant value as explained above.
b. Hemispheric barotropic 24-hr. and 48-hr. height forecasts from each input data deck were then produced by the current operational forecast code. This forecast procedure converts the input heights by using the balance equation (c. f. Shuman [6]) and produces the forecast internally from the resulting initial field of stream potential. The 24-hr. and

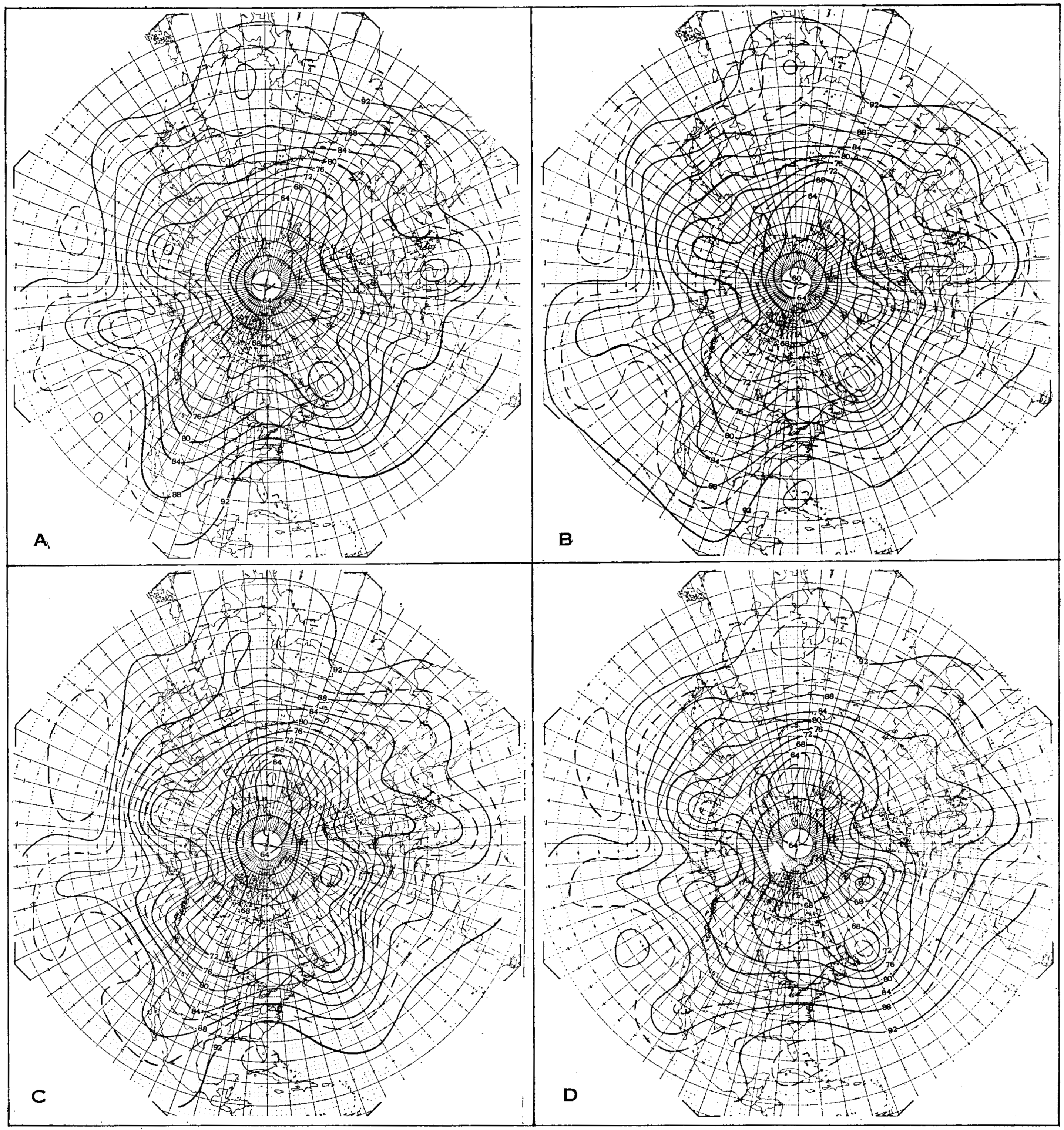

Figure 3.—48-hour forecasts made using the analyses of figure 2 as initial data.

48-hr. height output punch card decks are obtained by inverting back through the balance equation.

c. For each of the cases, the initial, 24-hr., and 48-hr. height field decks involving the data mesh size of figure 1A were compared with the corresponding decks involving the data mesh sizes of figure 1B, 1C, and 1D. A map showing height differences at grid points and a map of the corresponding wind errors were produced for each such comparison. Further, the wind errors were sorted by 10-knot intervals to indicate a frequency distribution.

Figure 3 displays the 48-hr. forecasts resulting from the corresponding initial charts of figure 2. Important phase differences are to be noted particularly in the

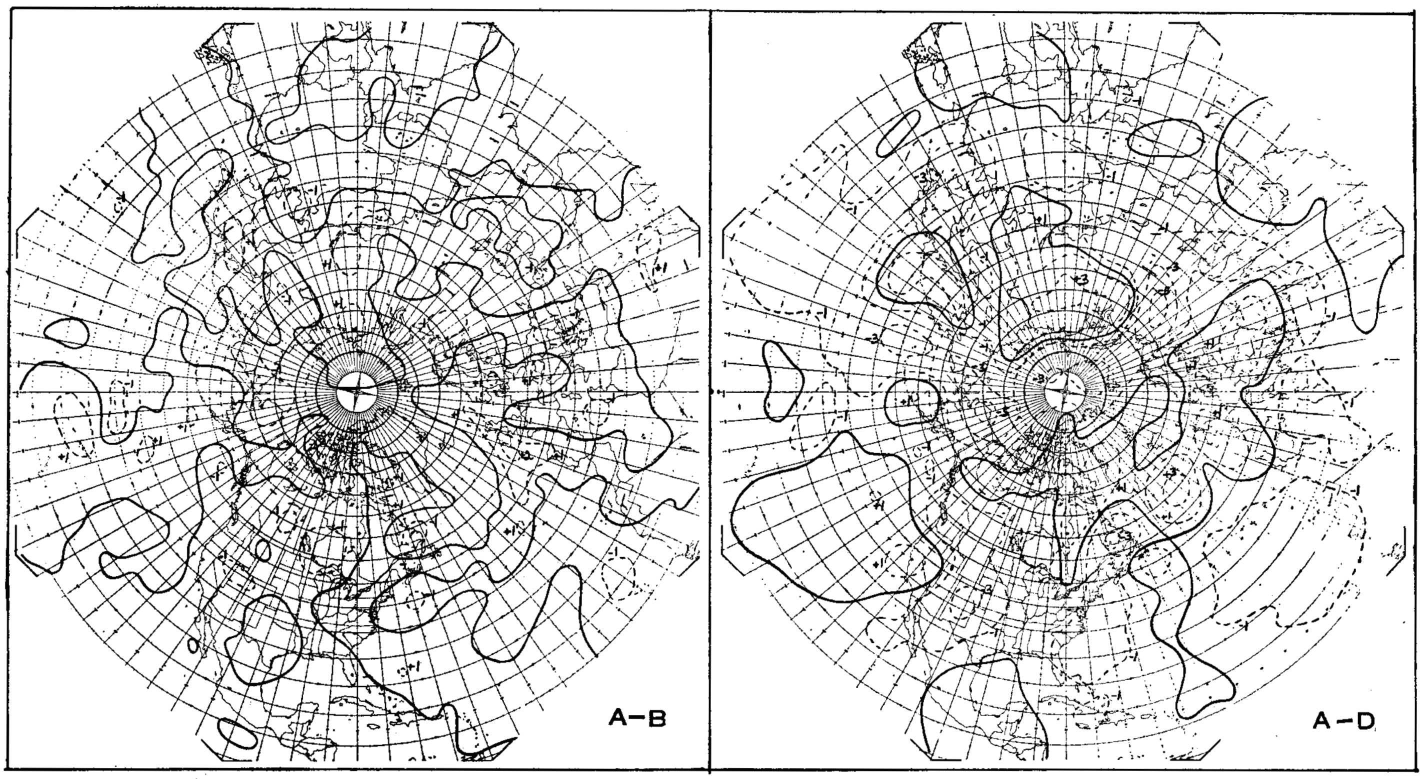

FIGURE 4.—Height error patterns (hundreds of feet) in Case 1 initial charts. (Left) Map of figure 2B compared with 2A. (Right) Map of figure 2D compared with 2A.

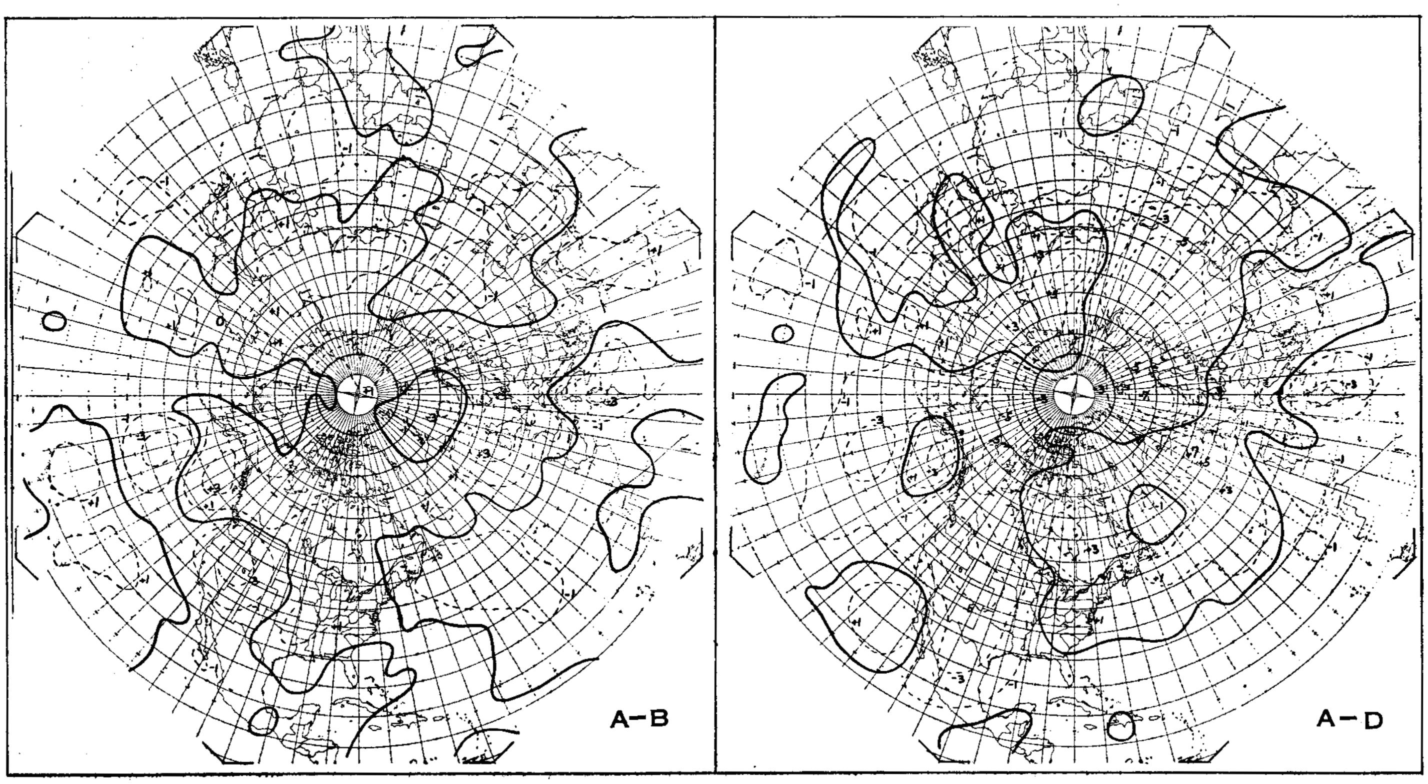

FIGURE 5.—Changes in the Case 1 height error fields at 48 hours (in hundreds of feet).

Table 1.—*Wind error frequency distribution: subjective analyses. (Percent of 1785 points with specified error.)*

Wind error (knots)	Initial			24-hr. forecast			48-hr. forecast		
	Max-Inter	Max-Min	Max-S-min	Max-Inter	Max-Min	Max-S-min	Max-Inter	Max-Min	Max-S-min
				CASE 1					
0-10	66.2	52.5	42.0	68.6	55.2	41.5	55.4	43.4	32.7
10-20	29.2	36.5	42.9	27.7	35.6	39.4	34.5	37.0	39.3
20-30	4.1	8.5	11.3	3.4	7.8	14.8	7.5	13.4	17.6
30-40	.4	1.8	3.0	.2	1.2	3.9	2.2	4.9	7.1
40-50	.1	.3	.7	-	.3	.7	-	1.1	2.2
50-60	-	.1	.2	-	-	.7	-	.3	1.0
60-70	-	.1	-	-	-	-	-	.4	.2
70-80	-	-	-	-	-	-	-	-	.0
80-90	-	-	-	-	-	-	-	-	.1
				CASE 2					
0-10	54.3	48.9	36.6	57.4	56.5	37.2	48.7	48.0	33.9
10-20	39.2	40.8	43.3	34.5	35.7	38.7	39.4	40.4	35.1
20-30	5.9	8.3	14.2	6.1	6.2	16.2	8.9	9.5	17.9
30-40	.6	1.5	4.3	1.3	1.3	5.5	1.7	1.8	7.0
40-50	-	.4	1.3	.4	.2	1.4	.9	.3	3.0
50-60	-	.2	.2	.1	-	.6	.3	-	1.6
60-70	-	-	-	.1	-	.2	.1	-	.8
70-80	-	-	-	-	-	.1	-	-	.2
80-90	-	-	-	-	-	.1	-	-	.1
90-100	-	-	-	-	-	.1	-	-	.1
100-110	-	-	-	-	-	-	-	-	.1
110-120	-	-	-	-	-	-	-	-	.1

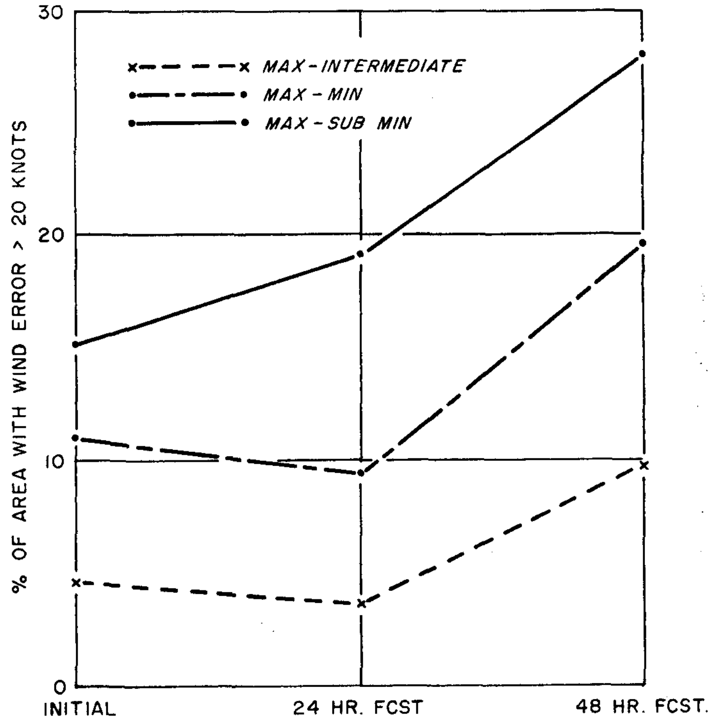

Figure 6.—Wind errors versus forecast range, Case 1, subjective analysis. Max=grid array A in figure 1, Intermdt=grid array B, Min=grid array C, and Sub-Min=grid array D.

trough feature in the region near California. Wind errors are even more striking. Figure 4 illustrates height error patterns for the comparison between initial charts having data coverages shown in figure 1A and 1B and similarly for coverages in figure 1A and 1D. In figure 4 the cellular pattern of errors for A–B is constrained to be small in scale as compared to similar quantities displayed for A–D. The changes in the error fields at 48 hours can be seen from the similar charts in figure 5.

The principal wind error results are given in table 1. Errors of each category are in terms of percent of total area computed on the basis of 1785 internal points. Wind errors were computed from the height error charts by considering the gradients diagonally across the grid squares of the basic 1977-point grid. This length corresponds to about 4° of latitude and approximates the portion of gradient normally used in hand measurements with a geostrophic wind scale. The results may be displayed graphically in a variety of ways. With such a limited sample it is perhaps better to suppress some of the detail in the frequency distribution of errors. A simple two-category breakdown is presented in figure 6. Percentage of area with wind errors over 20 knots is plotted as a function of forecast duration for the three comparisons involved. Even for rather intense winter regimes, a 20-knot wind error represents a large portion of the actual wind at 500 mb. From figure 6 it is first of all evident that the areas of large error increase rather uniformly as the data become more sparse with the sub-minimum (Pacific-type) coverage producing almost three times as much error as does the intermediate (Canadian-type) coverage. In addition, it is of interest to note that the minimum error is reached at 24 hours on the lower two curves whereas it occurs at the outset for the most sparse data array. According to Thompson [3], ". . . if the scale of the initial error field were appreciably less than that of the true fluctuations, the error might actually decrease for a while." The evidence here supports his analysis with the lower error curves actually diminishing at 24 hours. Further, the behavior of the upper curve, as relates to reality, implies that there are limited regions in the central Pacific, where, on occasion, predictability is at best extremely marginal. Case 2 produced a similar picture but with the lower curve much closer to the middle curve. Since the analysis with data corresponding to Canadian coverage was carried out by a different analyst one might speculate that different habits of analysis might be appearing. A comparison of the error charts for the two lower curves suggests rather that the analysis with Canadian coverage was smoothed in excess of the error tolerance in the data thereby eliminating some of the real small-scale error prescribed by the network. An average of the two cases is presented in figure 7.

The same evidence can be presented so as to emphasize more directly the question of data coverage. Figure 8 uses the same error parameter in the vertical coordinate and uses spacing between observations as horizontal coordinate. This treatment permits an estimate of error magnitude for any uniform data array. Unfortunately actual observations are not uniformly spaced except to a rough degree over limited areas. In some actual arrays one sees that isolated reports are called upon to yield detail that is clearly impossible. If one counts the number of observations in the outlined area of figure 9, excluding reconnaissance data, and weights all of them

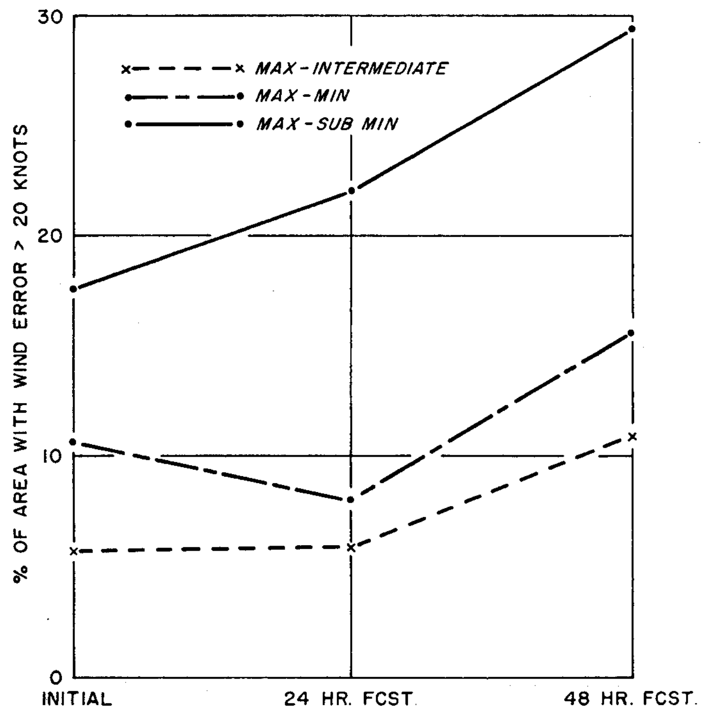

FIGURE 7.—Wind errors versus forecast range, average of Cases 1 and 2, subjective analysis.

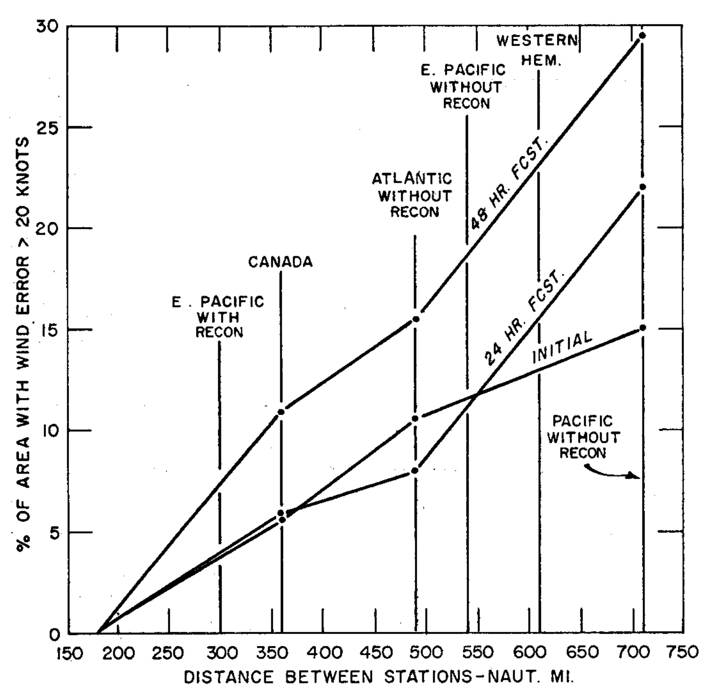

FIGURE 8.—Wind errors versus data density, Cases 1 and 2, subjective analysis. Same error parameter as in figure 7.

evenly, the mean distance between reports is about 540 n. miles. This is an extremely generous gesture since the analysis could obviously be improved by a more uniform redistribution. Even so, a uniform array of this dimension already is in a rather intolerable range of error as shown by figure 8. With the addition of the reconnaissance reports the again generous equivalent uniform array

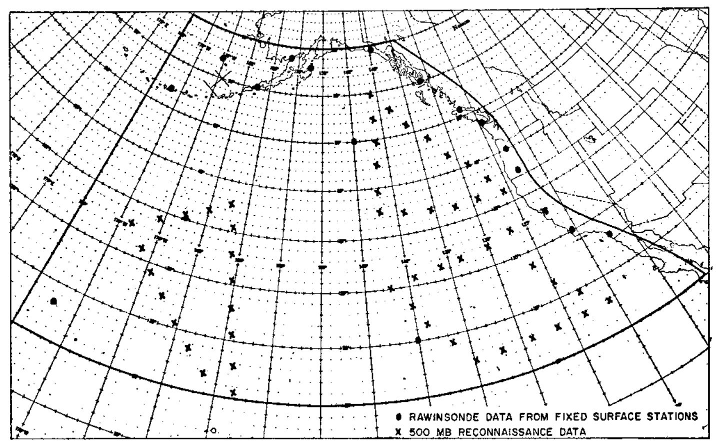

FIGURE 9.—Data coverage in eastern Pacific area.

has a data spacing of about 300 n. miles. The array, of course, still is not uniform and some of the reports are not strictly synoptic. By weighting reports by the area represented and including ". . . North America, the Caribbean Sea, the North Atlantic and most of the Pacific . . .", Thompson [3] arrived at an average distance of 610 n. miles as an equivalent mesh length for much of the Northern Hemisphere. The implication here is that if a certain data density is considered adequate for the resolution of synoptic features of a given scale, then more closely spaced data are partially redundant. Viewing from the standpoint of resolving the broad-scale features of the flow, Newton [1] cited several instances where a slight redistribution would greatly increase the value of reporting stations. The present study certainly supports such suggestions.

4. FURTHER TESTS USING OBJECTIVE ANALYSIS

The previously mentioned inhomogeneity in the subjective analyses of the second case and the large amount of chart work involved led to an attempt to enlarge the sample by using objective analysis. Accordingly a code was produced which formed an artificial input data tape to be used by the objective analysis program presently in routine use by the JNWP Unit. The data tape thus prepared contained data for the same arrays and modified them in the same manner as in the subjective analysis cases. One minor difference involved the application of random errors to the separate wind components thereby introducing variations in direction as well as speed.

The subjective analyst relies on data at other levels and a qualitative image from past data as a starting point. The technique of objective analysis used here relies more on a highly satisfactory quantitative image from past data—namely a 12-hour prognostic chart. In the present study, provision for a representative first approximation to the analyses posed a problem. In keeping with the spirit of the first cases, such a first approximation to the analysis should also reflect the analysis problems of the reduced data array. To avoid extensive

TABLE 2.—*Wind error frequency distribution: objective analyses. (Percent of 1221 points with specified error.)*

Wind error (knots)	Initial		24-hr. forecast		48-hr. forecast	
	Max-Min	Max-S-min	Max-Min	Max-S-min	Max-Min	Max-S-min
CASE 3						
0-10	56.6	42.8	58.4	43.2	56.8	37.8
10-20	33.7	39.1	33.4	38.2	34.8	38.3
20-30	7.4	13.7	6.9	13.2	7.0	15.6
30-40	1.6	3.2	1.2	3.4	1.2	5.5
40-50	.6	.8	.1	1.5	.2	2.0
50-60	.1	.4	----------	.3	----------	.7
60-70	----------	----------	----------	.1	----------	.1
CASE 4						
0-10	60.0	46.8	65.2	44.4	58.8	41.4
10-20	31.0	35.6	28.8	37.0	33.3	38.7
20-30	6.9	12.0	5.0	12.4	6.3	13.7
30-40	1.5	3.8	.9	4.9	1.6	4.5
40-50	.5	1.7	.1	1.3	----------	1.4

iteration by machine to accomplish this goal, an alternate simulation scheme was employed. The operationally available 24-hour forecast from the preceding day was substituted as a first approximation. This somewhat less perfect image of the current upper flow pattern is considered to be comparable to an indirect derivation from concurrent lower-level information insofar as placement of major features is concerned. Such a forecast chart still would be expected to contain details of shear and curvature not resolvable with the most sparse data array of the experimental coverage network. Accordingly considerable smoothing was applied before substituting the forecast chart as the simulated first approximation for the analysis.

The resulting wind error information analogous to that in table 1 is presented in table 2. Case 3 employed somewhat less smoothing than did case 4. In general the main results of the subjective analysis experiments are duplicated. The comparisons involving the simulated intermediate data array (Canadian-type coverage) were not carried out. Figures 10 and 11 are to be compared with figure 6. Here one sees the same pattern. Initial analysis errors from the sub-minimum array (Pacific) are roughly twice as large as are the errors involving the minimum (Atlantic) array. Also the errors are reduced in both cases at 24 hours for the lower curve whereas they are not reduced in the upper curves. One important difference seems to be that the error growth rate is not as large for the upper curves in the objective analysis cases. Finally it should be pointed out the degree of smoothing was deliberately fixed in a separate trial case so that the scale of the error pattern initially was predominantly controlled by the mesh size of the reduced data arrays. To this extent the objective analyses were guided by the subjective tests. However this does not alter the result that, once the first approximation procedure is fixed, we see the same pattern of behavior as in the subjective cases when the data array is changed.

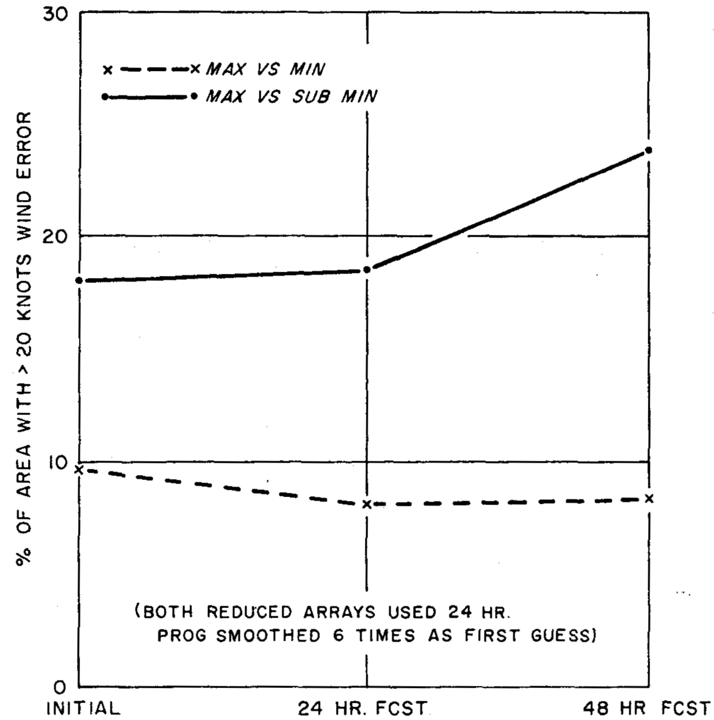

FIGURE 10.—Wind error versus forecast range, Case 3, objective analysis. Compare with figure 5.

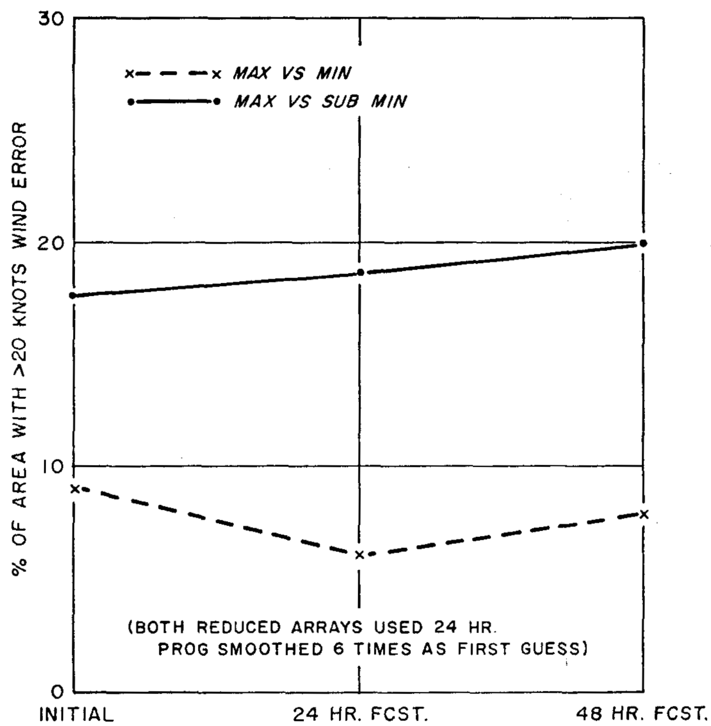

FIGURE 11.—Wind error versus forecast range, Case 4, objective analysis. Compare with figure 6.

5. SUMMARY AND CONCLUDING REMARKS

Two series of subjective analyses and the resulting barotropic forecasts indicate error behavior in accordance with that specified in Thompson's investigation. Specifically a measure of the forecast wind error is shown to be a function of forecast range, initial error, and scale. Further, such errors are shown to represent a very real practical problem in some areas for which forecasts are presently being issued. In such areas the future success of more sophisticated multi-parameter models may depend to a marked degree upon improvement in upper data coverage. Additional tests with objective analyses, which were somewhat dependent upon experience from the subjective analyses, suggest the same pattern of error behavior. Pending adoption of a baroclinic model for routine forecasts, additional tests of a similar character for the baroclinic case are planned. It seems intuitively obvious that the three-dimensional analysis problem, which requires proper phasing between pressure and temperature fields, places even greater requirements for adequate data coverage.

ACKNOWLEDGMENTS

The author is indebted to Lt. Col. P. D. Thompson for valuable consultations and to Mr. J. Murray Mitchell, Jr. for his assistance in chart analysis

REFERENCES

1. C. W. Newton, "Analysis and Data Problems in Relation to Numerical Prediction," *Bulletin of the American Meteorological Society*, vol. 35, No. 7, Sept. 1954, pp. 287–294.
2. W. H. Best, "Differences in Numerical Prognoses Resulting from Differences in Analyses," *Tellus*, vol. 8, No. 3, Aug. 1956, pp. 351–356.
3. P. D. Thompson, "Uncertainty of Initial State as a Factor in the Predictability of Large Scale Atmospheric Flow Patterns," *Tellus*, vol. 9, No. 3, Aug. 1957, pp. 275–295.
4. G. 'Arnason, Personal Communication, 1958.
5. R. Berggren, "On the Accuracy of 500 mb. Analysis with Special Reference to Numerical Forecasting," *Tellus*, vol. 9, No. 3, Aug. 1957, pp. 323–340.
6. F. G. Shuman, "Numerical Methods in Weather Prediction: I. The Balance Equation," *Monthly Weather Review*, vol. 85, No. 10, Oct. 1957, pp. 329–332.

Weather Note

WORLD RECORD LOW TEMPERATURE

Antarctica has again broken the world low temperature record. A message received from Morton J. Rubin at Mirny, Antarctica announced the following minimum temperature readings: At Vostok (78°27′ S., 106°52′ E.) −85.8° C. (−122.4° F.) was measured between 1200 GMT August 7 and 0000 GMT August 8, 1958; at Sovietskaya (78°24′ S., 87°35′ E.) −86.7° C. (−124.1° F.) between 1900 and 2000 GMT, on August 9, 1958. The Sovietskaya minimum was measured with a remote electrical resistance thermometer with a known correction, exposed in the shelter. "The temperature was lower immediately afterward but no calibration for thermometer," the message stated.

CORRECTION

MONTHLY WEATHER REVIEW, vol. 86, July 1958, p. 253: In the second equation in column one b_1 should read $\sqrt{b_1}$.